Convergence to a Convention in a 2x2 Coordination Game with Adaptive Learning

Ethan Holdahl

University of Oregon

September 7, 2022

Coordination Game G

Player B

		0	1
	0	$a_{00}, b_{00}{}^{*}$	a_{01}, b_{01}
Player A			
	1	a_{10}, b_{10}	$a_{11}, b_{11}{}^{*}$

Lemma 1 will be used in the proof of Theorem 1.

Let G be a 2x2 coordination game, and let $P^{m,s,\epsilon}$ be adaptive learning with memory m, sample size s, and error rate ϵ. Denote the players A and B with strategies $\{0, 1\}$ as depicted above where (0,0) and (1,1) constitute pure strategy Nash Equilibria. Define $h_{-i}^{t} = (x_{-i}^{t-m}, .., x_{-i}^{t-1})$ as the most recent m records of all players except player i at time t. Define R_{i}^{t} as the set of s records sampled by player i in period t from h_{-i}^{t}. Define BR_{i}^{t} as player i's best response to R_{i}^{t}.

Lemma 1. Let G be a 2x2 coordination game, and let $P^{m,s,\epsilon}$ be adaptive learning with memory m, sample size s, and error rate ϵ. If at any period n $\exists$ a strategy profile $x^* = (x_1^*, x_2^*)$ that constitutes a Nash Equilibrium where $x_1^* \in \{BR_1^n\}$ and $x_2^* \in \{BR_2^n\}$ for players 1 and 2 then there exists a positive probability that each player i plays x_i^* as a best response for every period $e \geq n$.

Proof of Lemma 1. Assume at period n $\exists$ an action $x^* \in BR_A^n \cap BR_B^n$. I must show that there exists a positive probability that each player plays x^* as a best reply for every period $e \geq n$.

I use proof by induction.

Base Step: I show that if there exists an action x^* that is a best response for both players in period n then there exists a positive probability that both players play action x^* in period n.

Clearly, if x^* is a best response for both players then there exists a positive probability that both players play action x^* in period n.

Inductive Step: I show that $\forall e \geq n$ if action x^* was played as a best response for both players in period e then there exists a positive probability that action x^* is played as a best response in period $e + 1$.

Assume action x^* was played as a best response for both players in period e. I

must show that there exists a positive probability that action x^* is played as a best response in period $e + 1$.

Note that since $x_i^t \in \{0, 1\}$ $\forall t$, the proportion of times that player $j \neq i$ played 1 in R_i^t, the set of s records in period t, is simply $\sum_{r \in R_i^t} \frac{r}{s}$. For example, if $R_i^t = (0, 1, 1, 1, 0, 0, 1, 1)$ then then proportion of times player $j \neq i$ played 1 in player i's sample, $R_i^t = \sum_{r \in R_i^t} \frac{r}{s} = \frac{5}{8}$.

Without loss of generality assume that $x^* = 1$. That means action 1 is a best response for each player $i \in \{A, B\}$ to R_i^e, the set of s records sampled by player i in period e.

Let $\alpha_i \in (0, 1)$ be the smallest probability that Player $j \neq i$ plays action 1 such that Player i's best response is playing action 1.

That means that for $i \in A, B$:

$$(1) \sum_{r \in R_i^e} \frac{r}{s} \geq \alpha_i$$

Now consider the set of s records sampled by player i in $e + 1$: R_i^{e+1}.

In period $e + 1$ each player i samples s records from h_{-i}^{e+1}, the most recent m records of all players except player i at time t. Note that $|h_{-i}^e \cap (h_{-i}^{e+1})'| = 1$, that is to say that there is only 1 record in h_{-i}^e that is not in h_{-i}^{e+1}. Since $R_i^e \subseteq h_{-i}^e$ I know that $|R_i^e \cap (h_{-i}^{e+1})'| \leq 1$. That means there is at most 1

3

record in player i's sample in period e that is not able to be sampled in period $e + 1$. This means that there is a positive probability that in period $e + 1$, each player i samples $s - 1$ records from the set $h_{-i}^{e+1} \cap R_i^e$ and the most recent record, x_{-i}^e. Assume both players samples in period $e + 1$ fit this criteria and define $c_{-i}^e = R_i^e \backslash R_i^{e+1}$, the record that was in the sample in period e but not in period $e + 1$ for player i.

So, $c_{-i}^e = R_i^e \backslash R_i^{e+1}$ and $x_{-i}^e = R_i^{e+1} \backslash R_i^e$. Consequently, I know that for each player i:

$$\frac{c_{-i}^e}{s} + \sum_{r \in R_i^{e+1}} \frac{r}{s} = \frac{x_{-i}^e}{s} + \sum_{r \in R_i^e} \frac{r}{s}$$

However, I know that $x_{-i}^e = 1$ and $c_{-i}^e \in \{0, 1\}$. So, $x_{-i}^e \geq c_{-i}^e$. Adding to both sides I get:

$$\frac{c_{-i}^e}{s} + \frac{x_{-i}^e}{s} + \sum_{r \in R_i^{e+1}} \frac{r}{s} \geq \frac{x_{-i}^e}{s} + \frac{c_{-i}^e}{s} + \sum_{r \in R_i^e} \frac{r}{s}$$

So

$$(2) \quad \sum_{r \in R_i^{e+1}} \frac{r}{s} \geq \sum_{r \in R_i^e} \frac{r}{s}$$

Using (1) and (2) by transitivity I get:

$$\sum_{r \in R_i^{e+1}} \frac{r}{s} \geq \sum_{r \in R_i^e} \frac{r}{s} \geq \alpha_i$$

So for each player $i \in \{A, B\}$ action 1 is a best response to the sample R_i^{e+1}.

Since action 1 is a best response for both players in period $e + 1$ there exists a positive probability that each plays action 1 in period $e + 1$.

Since both the base case and the inductive step has been shown, by mathematical induction I have proven that if both players play x_i^* as a best response in period n, then there exists a positive probability that both players play x_i^* as a best reply for every period $e \geq n$. $\square$

Theorem 1. Let G be a 2x2 coordination game where the possible actions for both players A and B are $\{0, 1\}$, and let $P^{m,s,\epsilon}$ be adaptive learning with memory m, sample size s, and error rate ϵ. Let $\alpha \in (0, 1)$ be the smallest probability that Player B plays action 1 such that Player A's best response is playing action 1. Likewise, Let $\beta \in (0, 1)$ be the smallest probability that Player A plays action 1 such that Player B's best response is playing action 1.

If $s < m$ or $\lceil \alpha s \rceil \neq \lceil (1 - \beta)s \rceil$ then from any initial state, the unperturbed process $P^{m,s,0}$ converges with probability one to a convention and locks in.

Proof of Theorem 1. Assume both players only sample the most recent s records. If at any period, t, both players share a best response, x^*, then by Lemma 1, there exists a positive probability that both players play action x^* as a best reply for each period $e \geq t$. If this happens then after period $t + m - 1$ the entire memory is filled with both players playing action x^*. Since $\epsilon = 0$ and since both players could then only sample records of the other player playing x^*, both would continue to play x^* as a best response for every period thereafter.

So, it remains to be proved that a convention can be reached when both players would never share a best response when sampling the most recent s records, or, that by sampling the most recent s records that there would

6

always exist a period t where both players share a best response.

First, I will show that when $\lceil \alpha s \rceil \neq \lceil (1 - \beta)s \rceil$, that by sampling the most recent s records there will always exist a period t where both players share a best response.

I use proof by contradiction. Assume that when sampling the most recent s records both players never share a best response and that $\lceil \alpha s \rceil \neq \lceil (1 - \beta)s \rceil$. Because this is adaptive learning in a coordination game, we know that if players aren't playing a Nash Equilibria that eventually one or both of their best responses will change to the strategy that the other player was playing. In order for players to never share a best response, it must be the case that whenever either players' best response changes, the best response of the other player simultaneously changes as well. Given enough periods, this pattern of simultaneously switching actions will yield histories of play for each player that are exactly the inverse of the history of play of the other player. In the context of this game where the possible actions for both players are $\{0, 1\}$ and both players best responses change at the same time when sampling the most recent s records, this means that the amount of records in player A's sample necessary for player A to play action 1 as a best response is equal to the amount of records in player B's sample necessary for player B to play action 0 as a best response: $\lceil \alpha s \rceil = \lceil (1 - \beta)s \rceil$. However, this contradicts the

7

assumption that $\lceil \alpha s \rceil \neq \lceil (1 - \beta)s \rceil$. So, when sampling the most recent s records, it is impossible for both players to never share a best response if $\lceil \alpha s \rceil \neq \lceil (1 - \beta)s \rceil$.

So, I have shown that when $\lceil \alpha s \rceil \neq \lceil (1 - \beta)s \rceil$, if players only sampling the most recent s records there must exist a period t where both players share a best response.

Next, I will show that when $s < m$ if players never share a best response when sampling the most recent s records that a convention can still be reached with positive probability.

Assume that if players sampled the most recent s records that they would never share a best response. As discussed above, this requires both players to simultaneously switch their best responses to what the other player was previously playing. This is the case in example 2 in Young [1993]. Let t be a period where both players would simultaneously switch their best response if they sampled the most recent s records. In this period, let player A sample the most recent s records and, since $m > s$, we can let player B sample the most recent $s + 1$ records excluding the most recent record played by player A. Consequently, player A's best response changed to match player B's action in the previous period, but because player B's sampled the same set of records they did previously, their response did not change in period t. Call x^* the

action both players played as a best response in period t. So, by Lemma 1, there exists a positive probability that both players play action x^* as a best reply for each period $e \geq t$. If this happens then after period $t + m - 1$ the entire memory is filled with both players playing action x^*. Since $\epsilon = 0$ and since both players could then only sample records of the other player playing x^*, both would continue to play x^* as a best response for every period thereafter.

Thus, I have shown that when $s < m$ if players never share a best response when sampling the most recent s records that a convention can still be reached with positive probability.

I have proven that from any initial state when $s < m$ or $\lceil \alpha s \rceil \neq \lceil (1 - \beta)s \rceil$ and $\epsilon = 0$ a convention can be reached with positive probability and lock in. Since a convention can be reached from any arbitrary state and since conventions are absorbing states we know that as $T \to \infty$ that h^T, the state at time T converges with probability one to a convention in and locks in. ∎

References

H. P. Young. The evolution of conventions. *Econometrica: Journal of the Econometric Society*, pages 57–84, 1993.